

# Positivity, Discontinuity, Finite Resources and Nonzero Error for Arbitrarily Varying Quantum Channels

H. Boche, J. Nötzel

Electronic addresses: {boche, janis.noetzel}@tum.de

Lehrstuhl für Theoretische Informationstechnik, Technische Universität München,  
80290 München, Germany

May 8, 2014

## Abstract

This work is motivated by a quite general question: Under which circumstances are the capacities of information transmission systems continuous? The research is explicitly carried out on arbitrarily varying quantum channels (AVQCs).

We give an explicit example that answers the recent question whether the transmission of messages over AVQCs can benefit from distribution of randomness between the legitimate sender and receiver in the affirmative.

The specific class of channels introduced in that example is then extended to show that the deterministic capacity does have discontinuity points, while that behaviour is, at the same time, not generic: We show that it is continuous around its positivity points. This is in stark contrast to the randomness-assisted capacity, which is always continuous in the channel. Our results imply that the deterministic message transmission capacity of an AVQC can be discontinuous only in points where it is zero, while the randomness assisted capacity is nonzero.

Apart from the zero-error capacities, this is the first result that shows a discontinuity of a capacity for a large class of quantum channels. The continuity of the respective capacity for memoryless quantum channels had, among others, been listed as an open problem on the problem page [29] of the ITP Hannover for about six years before it was proven to be continuous.

We also quantify the interplay between the distribution of finite amounts of randomness between the legitimate sender and receiver, the (nonzero) decoding error with respect to the average error criterion that can be achieved over a finite number of channel uses and the number of messages that can be sent. This part of our results also applies to entanglement- and strong subspace transmission. In addition, we give a new sufficient criterion for the entanglement transmission capacity with randomness assistance to vanish.

## I Introduction

A key property of a communication system is a continuous dependence of its performance on the system parameters. If small perturbations of the system lead to dramatic losses in the performance, it will most likely not be too widely used and instead be replaced by a more robust system.

The most fundamental property of a communication system is its capacity, and the very definition of capacities in quantum Shannon information theory is generically such that a straightforward application of Fekete's Lemma proves their existence as a real number.

One could now raise the question whether a similarly straightforward method exists that enables one to prove the continuity of channel capacities. The known methods are built upon the continuity of the entropy and require an explicit entropic formula for the capacity. Using them, one can prove [27] that the capacities of memoryless quantum channels are continuous, and these methods carry over to the randomness-assisted capacities of AVQCs [7]. At the same time, the zero-error capacities (see e.g. [20])

are certainly not continuous (this has e.g. been proven in [7]), and there exists no explicit formula for computing them.

At this point, we observe that there is no known formula for the deterministic message transmission capacity of an AVQC, so we may rightfully ask whether it is continuous or not. Our investigation revealed that there do indeed exist discontinuity points for the unassisted (deterministic) message transmission capacity of an AVQC. This finding is intimately connected to the difference between randomness-assisted and unassisted message transmission capacity of AVQCs, a difference which has explicitly been conjectured in [17] and is proven in this paper.

The question whether a general method for proving continuity of a communication system exists remains open.

As an additional result we give bounds on the functional dependence between blocklength, error and the amount of common randomness needed to achieve that error.

In the following, we will explain the model of an arbitrarily varying channel and provide examples for communication scenarios whose essential features are captured by the model. We then explain the effect of shared randomness for these systems and state a corresponding result. In close connection, we discuss the relevance of continuity of capacities, state results and give examples. Finally, we quantify the interplay between finite errors, block length and the amount of common randomness needed to achieve that error.

Imagine a situation where a sender wants to transmit for example messages to a remote receiver. They each have access to a quantum system which is modeled on a finite dimensional Hilbert space and are connected by a quantum channel. Dependent on the message he wants to transmit the sender prepares some quantum state, which is then transmitted to the receiver over the channel. The question then is, whether the receiver can infer which message the sender intended to send just by performing measurements on the output states.

We assume that multiple channel uses are available and that the channel does not have a memory - but instead assume the existence of a jammer, which tries to prevent the two legal parties from communicating properly. Such a situation can arise e.g. in secret key distribution or transmission scenarios over quantum channels as developed by Devetak in [19], but when the evil third party is either not interested in or unable to do eavesdropping on the legal communication, but has some influence on the channel between the legal parties. The power of the jammer is, in the model chosen here, precisely quantified by his ability to influence the channel:

He is able to choose, for each of the multiple channel uses, one out of a fixed set  $\mathfrak{J}$  of channels. This set is known to all three parties. The goal of sender and receiver is now to find encoding-and decoding procedures such that they can reliably transmit their data, no matter which choice the jammer makes. It can even be assumed that the jammer knows *in advance* how the encoding-decoding procedure of sender and receiver works. This assumption will always be satisfied in commercial communication systems, where standardized protocols are being used. The model that we just introduced is called an arbitrarily varying quantum channel. Note that, throughout the entire manuscript, we restrict attention to finite AVQCs, e.g. those for which  $|\mathfrak{J}| < \infty$  holds. The main reason for this is that it greatly simplifies proofs and puts a clean focus on the most relevant features of the systems under consideration.

Of course, the very same model can be formulated by using as the basic channels either classical, classical-quantum or quantum-classical channels, and the underlying systems that the three parties act upon could be described by any kind of physical theory. Another possible change in the model would be to enable the jammer to use quantum inputs to the system. In this work, we will stick to the model we described first.

The situation described by the model can, in these days, be found in denial-of-service attacks. It is important to note that the quality of an arbitrarily varying classical channel can not only be described by entropic quantities, as is the case for stationary memoryless channels. It has rather been found that its capability to transmit any messages at all is completely characterized by so-called *symmetrizability conditions*.

Let us get into a bit more detail here. It has been proven, first in [2] for classical arbitrarily varying

channels, then in [5] for classical-quantum arbitrarily varying channels that these systems exhibit a dichotomic behaviour: the message-transmission capacity under average error criterion,  $\overline{C}_{\text{det}}$ , is either zero or equals an easily computable number, called the random capacity  $\overline{C}_{\text{random}}$ . The latter quantity is the amount of messages that can be sent with transmission error approaching zero, when the number of channel uses goes to infinity and sender and receiver share a sufficiently large amount of shared randomness (polynomially much common randomness, in the number of channel uses, is sufficient). It turned out later [18, 21] that those arbitrarily varying channels  $\mathfrak{W}$  for which  $\overline{C}_{\text{det}}(\mathfrak{W}) = 0$  holds are exactly characterized by so-called “symmetrizability” conditions.

The dichotomic behaviour has been proven to hold true for both entanglement and message transmission over AVQCs in [7]. Another result of the work [7] was that encoding-decoding schemes for entanglement transmission are also good for strong subspace transmission and vice versa. The later work [17] showed that this is also true for message transmission under average- and maximal error criterion. These results enable us to restrict our discussion to the average error criterion and entanglement transmission henceforth.

Despite these achievements, it remained an open question until now whether shared randomness really helps the transmission of messages over AVQCs, and the same question remained open for entanglement- and strong subspace transmission.

More precisely, it has been conjectured in [7] that shared randomness does not increase the entanglement transmission capacity of AVQCs and in [17] that there exist examples of AVQCs  $\mathfrak{J}$  for which  $\overline{C}_{\text{random}}(\mathfrak{J}) > 0$  but  $\overline{C}_{\text{det}}(\mathfrak{J}) = 0$  holds.

In this work we provide exactly such an example.

We then study the continuity properties of  $\overline{C}_{\text{det}}$  for AVQCs. We find that  $\overline{C}_{\text{det}}$  is continuous around every AVQC  $\mathfrak{J}$  for which  $\overline{C}_{\text{det}}(\mathfrak{J}) > 0$  holds. Put into simple words: If a system which is modeled as an AVQC is ‘useful’ in the sense that  $\overline{C}_{\text{det}}(\mathfrak{J}) > 0$ , then this remains true even if small errors are present in the evaluation of the system parameters.

An obvious question that comes with the above two results is, whether there really exist discontinuities for the function  $\mathfrak{J} \mapsto \overline{C}_{\text{det}}(\mathfrak{J})$ . The continuity of the message- and entanglement transmission capacity of a stationary memoryless quantum channel has been an open problem for quite a while, it was posed by M. Keyl and listed in the open problem page [29] of R. Werner’s group since 2003. After partial results, it was completely solved by Leung and Smith in [27] in 2009, and answered in the affirmative: Both message- and entanglement transmission capacity are continuous for stationary memoryless quantum channels.

Quite on the contrary, we prove in this work that the message transmission capacity of AVQCs *without* assistance by shared randomness is not continuous. We do so by explicit construction of an example. This is the first example of a discontinuous behaviour of a quantum capacity other than the zero-error capacities [20].

Our previous results clearly demonstrate the importance of shared randomness for AVQCs. In [7], R. Ahlswede, I. Bjelaković and the authors showed that already a small amount of common randomness is sufficient to ensure that transmission of messages is possible at rates arbitrarily close to  $\overline{C}_{\text{random}}$ . The same holds true for transmission of entanglement. Building on that and the work [6] of R. Ahlswede and N. Cai, the authors were able to show in [17] that already the use of arbitrarily small amounts of correlation yield the same result.

This demonstrates that shared randomness has two important effects for AVQCs: First, it boosts the capacity to the maximally possible value, and second it stabilizes the system with respect to small changes (the capacity function with assistance by either unlimited shared randomness, positive correlation or small amounts of common randomness is always continuous).

This gives a strong motivation to start a closer investigation of the exact interplay between the system parameters, the error of message transmission at a specific block length and the amount of randomness used for stabilization of the system. This investigation is carried out in the last part of the paper. We give bounds on the number of shared secret bits (common randomness)  $K$  needed to achieve some pre-given maximal error  $\lambda$  within  $L$  channel uses. Assuming that the AVQC under consideration has  $|\mathbf{S}|$  constituents, the scaling law is roughly  $K \leq \frac{\log |\mathbf{S}|}{E \cdot \lambda}$ , where  $E$  is the reliability function of the *compound channel*  $\text{conv}(\mathfrak{J})$  and  $L$  is more implicitly given through  $E$ , roughly scaling

as  $L(1 - \frac{1}{L} \log L) \approx -E \log(E/\lambda)$ . In case that the AVQC  $\mathfrak{J}$  is symmetrizable, we note that the results of [7] imply  $\frac{1}{2\lambda} \leq K$ , and for non-symmetrizable AVQCs we know that  $K = 0$  is sufficient by the quantum-Ahlsvede dichotomy proven in [7].

Another important observation is that the number  $K$  of random bits needed to guarantee a certain quality of transmission is essentially independent of the number  $l$  of channel uses, if only  $l \geq L$  holds, and is indefinite for  $l < L$ .

The technique of proof we utilize here applies to entanglement transmission as well.

It is clear that a similar result could be obtained by using only correlation to first establish enough common randomness and then use it with the above stated bounds, but the exact trade-off between  $\lambda$ ,  $L$  and the 'amount' of correlation remains unclear and we leave that question open for future work.

A historical overview concerning the history of arbitrarily varying channels (in both the classical and the quantum case) can be found in [7]. Among the more recent developments are [7], [14], [32] and [17].

## II Notation

All Hilbert spaces are assumed to have finite dimension and are over the field  $\mathbb{C}$ . The set of linear operators from  $\mathcal{H}$  to  $\mathcal{H}$  is denoted  $\mathcal{B}(\mathcal{H})$ . The adjoint of  $b \in \mathcal{B}(\mathcal{H})$  is marked by a star and written  $b^*$ .

$\mathcal{S}(\mathcal{H})$  is the set of states, i.e. positive semi-definite operators with trace (the trace function on  $\mathcal{B}(\mathcal{H})$  is written  $\text{tr}$ ) 1 acting on the Hilbert space  $\mathcal{H}$ . The maximally mixed state with only one eigenvalue  $\dim(\mathcal{H})$  in  $\mathcal{S}(\mathcal{H})$  is written  $\pi_{\mathcal{H}}$  or, if no confusion can arise, simply  $\pi$ . Pure states are given by projections onto one-dimensional subspaces. A vector  $x \in \mathcal{H}$  of unit length spanning such a subspace will therefore be referred to as a state vector, the corresponding state will be written  $|x\rangle\langle x|$ . For a finite set  $\mathbf{X}$  the notation  $\mathfrak{P}(\mathbf{X})$  is reserved for the set of probability distributions on  $\mathbf{X}$ , and  $|\mathbf{X}|$  denotes its cardinality. For any  $l \in \mathbb{N}$ , we define  $\mathbf{X}^l := \{(x_1, \dots, x_l) : x_i \in \mathbf{X} \forall i \in \{1, \dots, l\}\}$ , we also write  $x^l$  for the elements of  $\mathbf{X}^l$ .

The set of completely positive trace preserving (CPTP) maps (also called quantum channels) between the operator spaces  $\mathcal{B}(\mathcal{H})$  and  $\mathcal{B}(\mathcal{K})$  is denoted by  $\mathcal{C}(\mathcal{H}, \mathcal{K})$ .

Closely related is the set of classical-quantum channels (abbreviated here using the term 'cq-channels') with finite input alphabet  $\mathbf{Z}$  and output alphabet  $\mathcal{K}$ , that arises from  $\mathcal{C}(\mathcal{H}, \mathcal{K})$  by setting  $d = |\mathbf{Z}|$  and restricting the inputs to matrices that are diagonal in any specific basis. This set is denoted  $CQ(\mathbf{Z}, \mathcal{K})$ . Both  $\mathcal{C}(\mathcal{H}, \mathcal{K})$  and  $CQ(\mathbf{Z}, \mathcal{K})$  are convex subsets of vector spaces.

For any natural number  $N$ , we define  $[N]$  to be the shorthand for the set  $\{1, \dots, N\}$ .

Using the usual operator ordering symbols  $\leq$  and  $\geq$  on  $\mathcal{B}(\mathcal{H})$ , the set of measurements with  $N \in \mathbb{N}$  different outcomes is written

$$\mathcal{M}_N(\mathcal{H}) := \{\mathbf{D} : \mathbf{D} = (D_1, \dots, D_N) \wedge \sum_{i=1}^N D_i \leq \mathbb{1}_{\mathcal{H}} \wedge D_i \geq 0 \forall i \in [N]\}. \quad (1)$$

To every  $\mathbf{D} \in \mathcal{M}_N(\mathcal{H})$  there corresponds a unique operator defined by  $D_0 := \mathbb{1}_{\mathcal{H}} - \sum_{i=1}^N D_i$ . Throughout the paper, we will assume that  $D_0 = 0$  holds. This is possible in our scenario, since adding the element  $D_0$  to any of the other  $D_1, \dots, D_N$  does not decrease the performance of a given code.

The von Neumann entropy of a state  $\rho \in \mathcal{S}(\mathcal{H})$  is given by

$$S(\rho) := -\text{tr}(\rho \log \rho), \quad (2)$$

where  $\log(\cdot)$  denotes the base two logarithm which is used throughout the paper.

The Holevo information is for a given channel  $W \in CQ(\mathbf{X}, \mathcal{H})$  and input probability distribution  $p \in \mathfrak{P}(\mathbf{X})$  defined by

$$\chi(p, W) := S(\overline{W}) - \sum_{x \in \mathbf{X}} p(x) S(W(x)), \quad (3)$$

where  $\overline{W}$  is defined by  $\overline{W} := \sum_{x \in \mathbf{X}} p(x)W(x)$

Given a bipartite random variable  $(X, Y)$ , its mutual information  $I(X, Y)$  is given by  $I(X, Y) := H(X) + H(Y) - H(X, Y)$ , where  $H(\cdot)$  is the usual Shannon entropy. If  $p \in \mathfrak{P}(\mathbf{X})$  for some set  $\mathbf{X}$  satisfying  $|\mathbf{X}| = 2$ , its Shannon entropy is also written as  $h(p(x)) := H(p)$ , where  $x \in \mathbf{X}$  is arbitrary.

For  $\rho \in \mathcal{S}(\mathcal{H})$  and  $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$  the entanglement fidelity (which was defined in [31]) is given by

$$F_e(\rho, \mathcal{N}) := \langle \psi, (id_{\mathcal{B}(\mathcal{H})} \otimes \mathcal{N})(|\psi\rangle\langle\psi|)\psi \rangle, \quad (4)$$

with  $\psi \in \mathcal{H} \otimes \mathcal{H}$  being an arbitrary purification of the state  $\rho$ .

For a finite set  $\mathcal{W} = \{W_s\}_{s \in \mathbf{S}} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$  or  $\mathcal{W} = \{W_s\}_{s \in \mathbf{S}} \subset CQ(\mathbf{Z}, \mathcal{K})$  we denote its convex hull by  $\text{conv}(\mathcal{W})$ . In the cases considered here the following will be sufficient. For a set  $\mathcal{W} := \{W_s\}_{s \in \mathbf{S}}$

$$\text{conv}(\mathcal{W}) = \left\{ W_q : W_q = \sum_{s \in \mathbf{S}} q(s)W_s, \quad q \in \mathfrak{P}(\mathbf{S}) \right\}. \quad (5)$$

As a measure of distance between channels we use the diamond norm  $\|\cdot\|_\diamond$ , which is given by

$$\|\mathcal{N}\|_\diamond := \sup_{n \in \mathbb{N}} \max_{a \in \mathcal{B}(\mathbb{C}^n \otimes \mathcal{H}), \|a\|_1 = 1} \|(\text{id}_n \otimes \mathcal{N})(a)\|_1,$$

where  $\text{id}_n : \mathcal{B}(\mathbb{C}^n) \rightarrow \mathcal{B}(\mathbb{C}^n)$  is the identity channel, and  $\mathcal{N} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$  is any linear map, not necessarily completely positive. The merits of  $\|\cdot\|_\diamond$  are due to the following facts (cf. [26]). First,  $\|\mathcal{N}\|_\diamond = 1$  for all  $\mathcal{N} \in \mathcal{C}(\mathcal{H}, \mathcal{K})$ . Thus,  $\mathcal{C}(\mathcal{H}, \mathcal{K}) \subset S_\diamond$ , where  $S_\diamond$  denotes the unit sphere of the normed space  $(\mathcal{B}(\mathcal{B}(\mathcal{H}), \mathcal{B}(\mathcal{K})), \|\cdot\|_\diamond)$ . Moreover,  $\|\mathcal{N}_1 \otimes \mathcal{N}_2\|_\diamond = \|\mathcal{N}_1\|_\diamond \|\mathcal{N}_2\|_\diamond$  for arbitrary linear maps  $\mathcal{N}_1, \mathcal{N}_2 : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$ .

We further use the diamond norm to define the function  $D_\diamond(\cdot, \cdot)$  on  $\{(\mathfrak{I}, \mathfrak{J}') : \mathfrak{I}, \mathfrak{J}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K})\}$ , which is for  $\mathfrak{I}, \mathfrak{J}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$  given by

$$D_\diamond(\mathfrak{I}, \mathfrak{J}') := \max\left\{ \sup_{\mathcal{N} \in \mathfrak{I}} \inf_{\mathcal{N}' \in \mathfrak{J}'} \|\mathcal{N} - \mathcal{N}'\|_\diamond, \sup_{\mathcal{N}' \in \mathfrak{J}'} \inf_{\mathcal{N} \in \mathfrak{I}} \|\mathcal{N} - \mathcal{N}'\|_\diamond \right\}.$$

For  $\mathfrak{I} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$  let  $\tilde{\mathfrak{I}}$  denote the closure of  $\mathfrak{I}$  in  $\|\cdot\|_\diamond$ . Then  $D_\diamond$  defines a metric on  $\{(\mathfrak{I}, \mathfrak{J}') : \mathfrak{I}, \mathfrak{J}' \subset \mathcal{C}(\mathcal{H}, \mathcal{K}), \mathfrak{I} = \tilde{\mathfrak{I}}, \mathfrak{J}' = \tilde{\mathfrak{J}'}\}$  which is basically the Hausdorff distance induced by the diamond norm.

### III Definitions

For the rest of this subsection, let  $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}} \subset \mathcal{C}(\mathcal{H}, \mathcal{K})$  denote a finite set of channels and  $\mathcal{H}, \mathcal{K}$  some arbitrary but fixed finite dimensional Hilbert spaces over  $\mathbb{C}$ . Henceforth, we follow the convention from [7], using the term 'the AVQC  $\mathfrak{I}$ ' as a linguistic shorthand for the mathematical object  $(\{\mathcal{N}_{s^l}\}_{s^l \in \mathbf{S}^l})_{l \in \mathbb{N}}$ . Due to the close correspondence between arbitrarily varying and certain compound channels, we will sometimes also encounter the case that  $\mathfrak{I}$  stands for the compound channel  $(\{\mathcal{N}_q^{\otimes l}\}_{\mathcal{N}_q \in \text{conv}(\mathfrak{I})})_{l \in \mathbb{N}}$ . In those cases, this will be explicitly mentioned. We will now define the entanglement transmission capacities of an AVQC. Corresponding coding theorems can be found in [7].

**Definition 1.** An  $(l, k_l)$ -random entanglement transmission code for  $\mathfrak{I}$  is a probability measure  $\mu_l$  on  $(\mathcal{C}(\mathcal{F}_l, \mathcal{H}^{\otimes l}) \times \mathcal{C}(\mathcal{K}^{\otimes l}, \mathcal{F}'_l), \sigma_l)$ , where  $\mathcal{F}_l, \mathcal{F}'_l$  are Hilbert spaces,  $\dim \mathcal{F}_l = k_l$ ,  $\mathcal{F}_l \subset \mathcal{F}'_l$  and the sigma-algebra  $\sigma_l$  is chosen such that the function  $(\mathcal{P}_l, \mathcal{R}_l) \mapsto F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_l \circ \mathcal{N}_{s^l} \circ \mathcal{P}_l)$  is measurable w.r.t.  $\sigma_l$  for every  $s^l \in \mathbf{S}^l$ .

Moreover, we assume that  $\sigma_l$  contains all singleton sets. An example of such a sigma-algebra  $\sigma_l$  is given by the product of sigma-algebras of Borel sets induced on  $\mathcal{C}(\mathcal{F}_l, \mathcal{H})$  and  $\mathcal{C}(\mathcal{K}, \mathcal{F}'_l)$  by the standard topologies of the ambient spaces.

The error of the code is given by  $\varepsilon_l := 1 - \int d\mu_l(\mathcal{R}^l, \mathcal{P}^l) F_e(\pi_{\mathcal{F}_l}, \mathcal{R}_l \circ \mathcal{N}_{s^l} \circ \mathcal{P}_l)$ .

**Definition 2.** A non-negative number  $R$  is said to be an achievable entanglement transmission rate for the AVQC  $\mathfrak{I} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$  with random codes and error  $\lambda \in [0, 1]$  if there is a sequence of  $(l, k_l)$ -random entanglement transmission codes such that

1.  $\liminf_{l \rightarrow \infty} \frac{1}{l} \log k_l \geq R$  and
2.  $\liminf_{l \rightarrow \infty} \inf_{s^l \in \mathbf{S}^l} \int F_e(\pi_{\mathcal{F}_l}, \mathcal{R}^l \circ \mathcal{N}_{s^l} \circ \mathcal{P}^l) d\mu_l(\mathcal{P}^l, \mathcal{R}^l) \geq 1 - \lambda$ .

The random entanglement transmission capacity  $\mathcal{A}_{\text{random}}(\mathfrak{J}, \lambda)$  of  $\mathfrak{J}$  with error  $\lambda$  is defined by

$$\mathcal{A}_{\text{random}}(\mathfrak{J}, \lambda) := \sup \left\{ R \in \mathbb{R}_+ : \begin{array}{l} R \text{ is an achievable entanglement trans-} \\ \text{mission rate for } \mathfrak{J} \text{ with random codes and error } \lambda \end{array} \right\}. \quad (6)$$

**Remark 1.** The reader with some previous knowledge about the topic of finite errors will notice that this definition differs from the classical one used e.g. in [8]. Precisely speaking, if one would define capacities with finite errors in the spirit of [8] and use the symbol  $\tilde{\mathcal{A}}_{\text{random}}$  for those, then one would set  $\tilde{\mathcal{A}}_{\text{random}} := \lim_{n \rightarrow \infty} \frac{1}{n} \log \max \{k_l : \exists (l, k_l) - \text{code for entanglement transmission with } \lambda_l \leq \lambda\}$ . Since  $\lambda \rightarrow \tilde{\mathcal{A}}_{\text{random}}(\mathfrak{J}, \lambda)$  is monotone increasing on  $[0, 1]$ , the limits  $\lim_{\varepsilon \rightarrow 0} \tilde{\mathcal{A}}_{\text{random}}(\mathfrak{J}, \lambda + \varepsilon)$  exist for every  $\lambda \in [0, 1]$ . It is then clear that the equality  $\mathcal{A}_{\text{random}}(\mathfrak{J}, \lambda) = \lim_{\varepsilon \rightarrow 0} \tilde{\mathcal{A}}_{\text{random}}(\mathfrak{J}, \lambda + \varepsilon)$  holds for all  $\lambda \in [0, 1]$ . We thus see that  $\mathcal{A}_{\text{random}}$  is simply the right-regularized version of  $\tilde{\mathcal{A}}_{\text{random}}$ .

While  $\tilde{\mathcal{A}}_{\text{random}}$  might be a practically more relevant definition, it is clear that the two definitions can lead to a different value in capacity only at discontinuity points of  $\tilde{\mathcal{A}}_{\text{random}}$ . Since both functions are monotone increasing on the interval  $[0, 1]$ , the number of such points is countable by [33], Theorem 4.30.

Notably, at  $\lambda = 0$ , one gets  $\mathcal{A}_{\text{random}}(\mathfrak{J}, 0) = \mathcal{A}_{\text{random}}(\mathfrak{J})$  for 'the' random capacity  $\mathcal{A}_{\text{random}}$  of an AVQC according to Definition 2 in [7], while  $\tilde{\mathcal{A}}_{\text{random}}(\mathfrak{J}, 0)$  gives the randomness-assisted zero-error capacity of an AVQC.

This latter point makes our definition fit seamlessly with the previous work [7, 17] on AVQCs. At the same time, we do not encounter a dramatically different behaviour in most cases. The same reasoning applies to all the other capacities defined in this paper.

Having defined random codes and random code capacity for entanglement transmission we are in the position to introduce their deterministic counterparts: An  $(l, k_l)$ -code for entanglement transmission over  $\mathfrak{J}$  is an  $(l, k_l)$ -random code for  $\mathfrak{J}$  with  $\mu_l(\{(\mathcal{P}^l, \mathcal{R}^l)\}) = 1$  for some encoder-decoder pair  $(\mathcal{P}^l, \mathcal{R}^l)$  (This explains our requirement on  $\sigma_l$  to contain all singleton sets) and  $\mu_l(A) = 0$  for any  $A \in \sigma_l$  with  $(\mathcal{P}^l, \mathcal{R}^l) \notin A$ . We will refer to such measures as point measures in what follows.

**Definition 3.** A non-negative number  $R$  is a deterministically achievable entanglement transmission rate for the AVQC  $\mathfrak{J} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$  with error  $\lambda \in [0, 1]$  if it is achievable in the sense of Definition 2 for random codes with point measures  $\mu_l$  and error  $\lambda$ .

The deterministic entanglement transmission capacity  $\mathcal{A}_{\text{det}}(\mathfrak{J}, \lambda)$  of  $\mathfrak{J}$  with error  $\lambda$  is given by

$$\mathcal{A}_{\text{det}}(\mathfrak{J}, \lambda) := \sup \left\{ R \in \mathbb{R}_+ : \begin{array}{l} R \text{ is an achievable entanglement trans-} \\ \text{mission rate for } \mathfrak{J} \text{ with deterministic codes and error } \lambda \end{array} \right\}. \quad (7)$$

We now switch attention to message transmission. From the results in [17], we know that average error criterion and maximal error criterion lead to the same capacity for AVQCs. Strictly speaking, this is a consequence of two facts: First, it does not really make sense to restrict the encoding functions to pure signal states in the quantum case, and second the two criteria are equivalent in the classical case as well, if one allows randomized encodings (see Ahlswede's paper [2], theorems 2 and 3).

**Definition 4** (Codes for message transmission over an AVQC). Let  $l \in \mathbb{N}$ . A random code for message transmission over  $\mathfrak{J}$  is given by a probability measure  $\gamma_l$  on the set  $(CQ(M_l, \mathcal{H}^{\otimes l}) \times \mathcal{M}_{M_l}, \Sigma_l)$ , where  $\Sigma_l$  again denotes a  $\sigma$ -algebra containing all singleton sets. Such  $\sigma$ -algebras exist, by arguments similar to those given in [7] and [12]. A deterministic code is then given by a random code  $\gamma_l$ , where  $\gamma_l$  is a point (also called Dirac) measure.

**Definition 5** (Achievability). A nonnegative number  $R$  is called achievable with random codes with error

$\lambda$  under the average error criterion if there exists a sequence  $(\gamma_l)_{l \in \mathbb{N}}$  of random codes satisfying both

$$1) \quad \liminf_{l \rightarrow \infty} \min_{s^l \in \mathbf{S}^l} \int \frac{1}{M_l} \sum_{i=1}^{M_l} \text{tr}\{D_i \mathcal{N}_{s^l}(\mathcal{P}(i))\} d\gamma_l(\mathcal{P}, \mathbf{D}) \geq 1 - \lambda \quad (8)$$

$$2) \quad \limsup_{l \rightarrow \infty} \frac{1}{l} \log M_l \geq R. \quad (9)$$

If the sequence  $(\gamma_l)_{l \in \mathbb{N}}$  can be chosen to consist of point measures only, then  $R$  is called *achievable with deterministic codes under the average error criterion*.

**Definition 6** (Message transmission capacities of an AVQC). *The corresponding capacities of  $\mathfrak{J}$  are defined as*

$$\overline{\mathcal{C}}_{\text{det}}(\mathfrak{J}, \lambda) := \sup \left\{ R : \begin{array}{l} R \text{ is achievable with deterministic codes} \\ \text{under the average error criterion with error } \lambda \end{array} \right\}, \quad (10)$$

$$\overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0) := \sup \left\{ R : \begin{array}{l} R \text{ is achievable with random codes} \\ \text{under the average error criterion with error } \lambda \end{array} \right\}. \quad (11)$$

The above definitions enable us to restate the main result of [7] which connects the entanglement transmission capacities via the deterministic message transmission capacity  $\overline{\mathcal{C}}_{\text{det}}(\cdot, 0)$ .

**Theorem 1** (Quantum Ahlswede dichotomy, cf. Theorem 1 in [7]). *Let  $\mathfrak{J} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$  be a finite AVQC.*

1. *For  $\text{conv}(\mathfrak{J})$  given in (5) we have*

$$\mathcal{A}_{\text{random}}(\mathfrak{J}, 0) = \lim_{l \rightarrow \infty} \frac{1}{l} \max_{\rho \in \mathcal{S}(\mathcal{H}^{\otimes l})} \inf_{\mathcal{N} \in \text{conv}(\mathfrak{J})} I_c(\rho, \mathcal{N}^{\otimes l}). \quad (12)$$

2. *Either  $\overline{\mathcal{C}}_{\text{det}}(\mathfrak{J}, 0) = 0$  or else  $\mathcal{A}_{\text{det}}(\mathfrak{J}, 0) = \mathcal{A}_{\text{random}}(\mathfrak{J}, 0)$ .*

It was also proven in [7] that  $\overline{\mathcal{C}}_{\text{det}}(\mathfrak{J}, 0) = 0$  holds if and only if the AVQC is *symmetrizable* according to definition 39 in [7]. This definition reads as follows:

**Definition 7** (Cf. definition 39 in [7]). *Let  $\mathbf{S}$  be a finite set and  $\mathfrak{J} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$  an AVQC.*

1.  *$\mathfrak{J}$  is called  $l$ -symmetrizable,  $l \in \mathbb{N}$ , if for each finite set  $\{\rho_1, \dots, \rho_K\} \subset \mathcal{S}(\mathcal{H}^{\otimes l})$ ,  $K \in \mathbb{N}$ , there is a map  $p : \{\rho_1, \dots, \rho_K\} \rightarrow \mathfrak{P}(\mathbf{S}^l)$  such that for all  $i, j \in \{1, \dots, K\}$*

$$\sum_{s^l \in \mathbf{S}^l} p(\rho_i)(s^l) \mathcal{N}_{s^l}(\rho_j) = \sum_{s^l \in \mathbf{S}^l} p(\rho_j)(s^l) \mathcal{N}_{s^l}(\rho_i) \quad (13)$$

*holds.*

2. *We call  $\mathfrak{J}$  symmetrizable if it is  $l$ -symmetrizable for all  $l \in \mathbb{N}$ .*

As mentioned already, every AVQC  $\mathfrak{J}$  is intimately connected to the compound quantum channel  $\text{conv}(\mathfrak{J})$ : the capacities of the AVQC  $\mathfrak{J}$  are often given by the respective formulas for the corresponding compound quantum channels  $\text{conv}(\mathfrak{J})$ . This connection especially shows up in the proof and formulation of our theorem 6, where we encounter the reliability functions of compound quantum channels. In order to define these, we first define codes, achievability and corresponding capacities for compound quantum channels:

**Definition 8** (Codes, achievability and capacities for compound quantum channels). *Let  $l \in \mathbb{N}$ . A code  $\mathfrak{C}_l$  for message transmission over the compound channel  $\mathfrak{J}$  is given by a natural number  $M_l$ , an encoding  $\mathcal{P} : [M_l] \rightarrow \mathcal{S}(\mathcal{H}^{\otimes l})$  and a decoding  $\mathbf{D} \in \mathcal{M}_{M_l}$ . The error  $\varepsilon_l$  associated to the code is given by*

$$\varepsilon_l := 1 - \min_{s \in \mathbf{S}} \frac{1}{M_l} \sum_{i=1}^{M_l} \text{tr}\{D_i \mathcal{N}_s^{\otimes l}\}. \quad (14)$$

A nonnegative number  $R$  is called *achievable* for the compound channel  $\mathfrak{J}$  with error  $\lambda \in [0, 1]$  if there exists a sequence  $(\mathfrak{C}_l)_{l \in \mathbb{N}}$  of codes for  $\mathfrak{J}$  satisfying both  $\limsup_{l \rightarrow \infty} \varepsilon_l \leq \lambda$  and  $\limsup_{l \rightarrow \infty} \frac{1}{l} \log M_l \geq R$ .

The capacity of the compound channel  $\mathfrak{J}$  with error  $\lambda$  is defined as

$$\overline{C}_{\text{det}}^{\text{comp}}(\mathfrak{J}, \lambda) := \sup \left\{ R : \begin{array}{l} R \text{ is an achievable rate for the compound} \\ \text{channel } \mathfrak{J} \text{ under the average error criterion and with error } \lambda \end{array} \right\}. \quad (15)$$

By changing the reliability criterion from average to maximal error, we can define the corresponding capacity  $C_{\text{det}}^{\text{comp}}$ , and switching to entanglement- or strong subspace transmission defines the capacities  $Q$  (see [13]) and  $Q_s$  in the obvious way.

These definitions enable us now to define the corresponding reliability functions:

**Definition 9** (Reliability Functions). *The reliability functions  $\overline{E}_m, E_m, E_e, E_s : \mathcal{C}(\mathcal{H}, \mathcal{K}) \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$  are defined by*

$$\overline{E}_m(\mathfrak{J}, R) := \sup \left\{ E : \begin{array}{l} \text{There is } \varepsilon > 0 \text{ and } N \in \mathbb{N} \text{ such that for all } l \geq N \text{ there is a code for} \\ \text{message transmission over the compound quantum channel } \mathfrak{J} \text{ satisfying} \\ \frac{1}{l} \log(M_l) \geq R - \varepsilon \text{ and } \varepsilon_l \leq 2^{-l(E-\varepsilon)}, \text{ with } \varepsilon_l \text{ being the average error.} \end{array} \right\} \quad (16)$$

$$E_m(\mathfrak{J}, R) := \sup \left\{ E : \begin{array}{l} \text{There is } \varepsilon > 0 \text{ and } N \in \mathbb{N} \text{ such that for all } l \geq N \text{ there is a code for} \\ \text{message transmission over the compound quantum channel } \mathfrak{J} \text{ satisfying} \\ \frac{1}{l} \log(M_l) \geq R - \varepsilon \text{ and } \varepsilon_l \leq 2^{-l(E-\varepsilon)}, \text{ with } \varepsilon_l \text{ being the maximal error.} \end{array} \right\} \quad (17)$$

$$E_e(\mathfrak{J}, R) := \sup \left\{ E : \begin{array}{l} \text{There is } \varepsilon > 0 \text{ and } N \in \mathbb{N} \text{ such that for all } l \geq N \text{ there is a code for} \\ \text{entanglement transmission over the compound quantum channel } \mathfrak{J} \\ \text{satisfying } \frac{1}{l} \log(k_l) \geq R - \varepsilon \text{ and error } \varepsilon_l \leq 2^{-l(E-\varepsilon)}. \end{array} \right\} \quad (18)$$

$$E_s(\mathfrak{J}, R) := \sup \left\{ E : \begin{array}{l} \text{There is } \varepsilon > 0 \text{ and } N \in \mathbb{N} \text{ such that for all } l \geq N \text{ there is a code for} \\ \text{strong subspace transmission over the compound quantum channel } \mathfrak{J} \\ \text{satisfying } \frac{1}{l} \log(k_l) \geq R - \varepsilon \text{ and error } \varepsilon_l \leq 2^{-l(E-\varepsilon)}. \end{array} \right\} \quad (19)$$

That above defined functions actually yield nonzero, finite values is not trivial in itself. It can, however, be explicitly read off the results in [13] that  $E_e(\mathfrak{J}, R) > 0$  holds if  $R < Q(\mathfrak{J})$  (where  $Q$  denotes the entanglement transmission capacity of the compound channel  $\mathfrak{J}$  in the work [13]), and that  $E_m(\mathfrak{J}, R)$  can be larger than zero follows from [11].

## IV Main Results

We now list our main results. If not specified otherwise,  $\mathfrak{J}$  denotes a finite AVQC throughout the entire section.

**Theorem 2.** *Let  $\mathfrak{J}$  consist of entanglement breaking channels that have the special form  $\mathcal{N}_s(\rho) := \sum_{x \in \mathbf{X}} \text{tr}\{\rho M_x\} \rho_{s,x}$ ,  $s \in \mathbf{S}$ , for some finite set  $\mathbf{S}$  and POVM  $\{M_i\}_{i=1}^M$  on  $\mathcal{H}$ . The following two statements are true:*

1. *If there are probability distributions  $\{p_x\}_{x \in \mathbf{X}} \subset \mathfrak{P}(\mathbf{S})$  such that*

$$\sum_{s \in \mathbf{S}} p_{x'}(s) \rho_{s,x} = \sum_{s \in \mathbf{S}} p_x(s) \rho_{s,x'} \quad \forall x, x' \in \mathbf{X}, \quad (20)$$

*then it holds  $\overline{C}_{\text{det}}(\mathfrak{J}, 0) = 0$ .*

2. *There exists an example of an AVQC satisfying the above conditions which additionally has the property  $\overline{C}_{\text{random}}(\mathfrak{J}, 0) > 0$ .*



**Remark 2.** Let us make a note on the intuition behind it. The channel  $\mathfrak{J}$  is the concatenation of a stationary memoryless qc-channel (measurement)  $\mathfrak{W}_1$  and an arbitrarily varying cq-channel  $\mathfrak{W}_2$  given by the states  $\{\rho_{s,x}\}_{s,x}$ . This combination ensures that the channel itself is entanglement-breaking, whence its capacity has a one-shot formula and, even more important, it is  $l$ -symmetrizable for all  $l \in \mathbb{N}$  if and only if it is 1-symmetrizable.

Using entangled inputs as signal states for  $\mathfrak{J}$  results in mixtures of product states after the application of  $\mathfrak{W}_1$ , so  $\mathfrak{W}_2$  sees a randomized code. But since we allow mixed inputs, this is equivalent to using just a randomized code with separable inputs for  $\mathfrak{W}_2$ . But on the subset of separable states signal states, 1-symmetrizability is equivalent to  $l$ -symmetrizability for all  $l \in \mathbb{N}$ , so no such code can transmit even a single bit with asymptotically vanishing error. Therefore, the deterministic capacity of  $\mathfrak{W}$  has to be equal to zero.

**Remark 3.** It is clear that the conjectured statement “for all finite AVQCs, it holds that  $\mathcal{A}_{\det}(\mathfrak{J}, 0) = \mathcal{A}_{\text{random}}(\mathfrak{J}, 0)$ ” is equivalent to saying that symmetrizability of a finite AVQC  $\mathfrak{J}$  according to Definition 39 in [7] implies that  $\mathcal{A}_{\text{random}}(\mathfrak{J}, 0) = 0$ .

It is also clear that either one of the above would imply that  $\mathcal{A}_{\det}$  is continuous, since  $\mathcal{A}_{\text{random}}$  is.

**Theorem 3.** [Discontinuity of  $\overline{\mathcal{C}}_{\det}(\cdot, 0)$ ] The capacity function  $\overline{\mathcal{C}}_{\det}(\cdot, 0) : \mathcal{C}(\mathcal{H}, \mathcal{K}) \rightarrow \mathbb{R}_+$  is not continuous.

More precisely, let  $\mathbb{C}^2 = \text{span}(\{e_1, e_2\})$  be naturally be embedded into  $\mathbb{C}^3 = \text{span}(\{e_1, e_2, e_3\})$ . Let the channel  $\mathcal{D}_\eta \in \mathcal{C}(\mathbb{C}^2, \mathbb{C}^3)$  be defined through  $\mathcal{D}_\eta(X) := (1 - \eta)X + \eta \cdot \text{tr}\{X\} \cdot \pi \ \forall X \in \mathcal{B}(\mathbb{C}^2)$ . For any  $\eta \in [0, 1]$  the sequence  $\mathfrak{J}_\lambda^\eta = \{\hat{\mathcal{N}}_{s,\eta,\lambda}\}_{s \in \mathbf{S}}$  of AVQCs defined by  $\hat{\mathcal{N}}_{s,\eta,\lambda} := (1 - \lambda)\mathcal{D}_\eta + \lambda\mathcal{N}_s$  with  $\{\mathcal{N}_s\}_{s \in \mathbf{S}}$  being the same set of channels as those constructed in Theorem 2 satisfies

$$\lim_{\lambda \rightarrow 1} \overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}_\lambda^\eta) = \overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0) \geq 0.5, \quad \overline{\mathcal{C}}_{\det}(\mathfrak{J}_\lambda^\eta, 0) = \overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0) \ \forall \lambda \in [0, 1], \quad \overline{\mathcal{C}}_{\det}(\mathfrak{J}_1) = 0. \quad (21)$$

On the other hand  $\lim_{\lambda \rightarrow 1} D_\diamond(\mathfrak{J}_\lambda^\eta, \mathfrak{J}) = 0$  for all  $\eta \in [0, 1]$ , so that  $\overline{\mathcal{C}}_{\det}(\cdot, 0)$  is discontinuous in the point  $\mathfrak{J}_1$ .

**Remark 4.** This is a first example of discontinuous behaviour of a quantum capacity other than the zero-error capacities. It is not clear to the authors yet, whether similar results could be proven for purely classical systems.

The example also highlights the stabilizing effect that is achieved by distribution of shared randomness in a communication system.

**Theorem 4** (Positivity of  $\overline{\mathcal{C}}_{\det}$  is stable). Let  $\mathfrak{J}$  be a finite AVQC satisfying  $\overline{\mathcal{C}}_{\det}(\mathfrak{J}, 0) > 0$ . There exists  $\delta_0 > 0$  such that for all finite AVQCs  $\mathfrak{J}'$  satisfying  $D_\diamond(\mathfrak{J}, \mathfrak{J}') \leq \delta_0$  it holds  $\overline{\mathcal{C}}_{\det}(\mathfrak{J}') > 0$ .

**Remark 5.** This result should be compared to the behaviour of the zero-error capacities, which are generically unstable (discontinuous at every point where they have a positive capacity). This comparison shows that communication systems exhibit a wide range of behaviour: Among them are continuous, discontinuous, stable and unstable ones.

This theorem has the following two corollaries:

**Corollary 1.** Let  $\mathfrak{J}$  be a finite AVQC and  $\overline{\mathcal{C}}_{\det}(\mathfrak{J}, 0) > 0$ . If  $\lim_{l \rightarrow \infty} D_\diamond(\mathfrak{J}, \mathfrak{J}_l) = 0$  for some sequence  $(\mathfrak{J}_l)_{l \in \mathbb{N}}$  of finite AVQCs, then also  $\lim_{l \rightarrow \infty} \overline{\mathcal{C}}_{\det}(\mathfrak{J}_l) = \overline{\mathcal{C}}_{\det}(\mathfrak{J}, 0)$ .

In other words,  $\overline{\mathcal{C}}_{\det}$  is lower-semi continuous, when restricted to (sequences of) finite AVQCs.

**Remark 6.** Of course, semi-continuity is a rather weak property (compared to continuity). In the history of quantum channel coding, lower semi-continuity of the entanglement transmission capacity of memoryless quantum channels has first been proven by [24] in 2002. The first complete proof of continuity of the (entanglement transmission) capacity for a one-parameter family of quantum channels (erasure channels) was given in [10] in 1998. Research on that line culminated in a proof by Leung and Smith [27] in 2009, showing that the message- and entanglement transmission capacities of a memoryless quantum channel are continuous. Their result also holds in the presence of a wiretapper.

The results of Leung and Smith easily extend to the case of AVQCs: The capacity  $\overline{C}_{\text{random}}$ , given by the formula in Lemma 2, is continuous. A method of proof can be picked up in [7], where a proof is given for the entanglement transmission capacity of an AVQC.

**Corollary 2.** Let  $\mathfrak{J}$  be a finite AVQC and  $\overline{C}_{\text{det}}(\mathfrak{J}, 0) > 0$ . If  $\lim_{l \rightarrow \infty} D_{\diamond}(\mathfrak{J}, \mathfrak{J}_l) = 0$  for some sequence  $(\mathfrak{J}_l)_{l \in \mathbb{N}}$  of finite AVQCs, then we have that

$$\mathcal{A}_{\text{det}}(\mathfrak{J}, 0) = \mathcal{A}_{\text{random}}(\mathfrak{J}, 0) = \lim_{l \rightarrow \infty} \mathcal{A}_{\text{det}}(\mathfrak{J}_l, 0). \quad (22)$$

Since by Theorem 3 we know that  $\overline{C}_{\text{det}}(\cdot, 0)$  is not continuous, it makes sense to characterize the points of discontinuity:

**Theorem 5** (Characterization of points of discontinuity). *Let  $\mathfrak{J}$  be an AVQC. The capacity function  $\overline{C}_{\text{det}}(\cdot, 0)$  is discontinuous in the point  $\mathfrak{J}$  if and only if  $\overline{C}_{\text{det}}(\mathfrak{J}, 0) = 0$ ,  $\overline{C}_{\text{random}}(\mathfrak{J}, 0) > 0$  and for every  $\varepsilon > 0$  there exists a finite AVQC  $\tilde{\mathfrak{J}}$  such that  $D_{\diamond}(\mathfrak{J}, \tilde{\mathfrak{J}}) < \varepsilon$  and  $\overline{C}_{\text{det}}(\tilde{\mathfrak{J}}, 0) > 0$ .*

Our forthcoming Definition 10 in Subsection V.2 that is needed in the proof of Theorem 4 motivated the formulation of the following observation:

**Lemma 1.** Let  $\mathfrak{J} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$  be a finite AVQC. If there exist  $l \in \mathbb{N}$  and functions  $p, q : \mathcal{S}(\mathcal{H}^{\otimes l}) \rightarrow \mathcal{P}(\mathbf{S}^l)$  such that for all  $\rho, \sigma \in \mathcal{S}(\mathcal{H}^{\otimes l})$  the equality

$$\sum_{s^l \in \mathbf{S}^l} p(\sigma)(s^l) \mathcal{N}_{s^l}(\rho) = \sum_{s^l \in \mathbf{S}^l} q(\rho)(s^l) \mathcal{N}_{s^l}(\sigma) \quad (23)$$

holds, then  $\text{conv}(\{\mathcal{N}_{s^l}\}_{s^l \in \mathbf{S}^l})$  contains an entanglement breaking channel and, consequently,  $\mathcal{A}_{\text{rand}}(\mathfrak{J}, 0) = 0$ .

**Theorem 6** (Random Code Reduction: finite error, finite randomness). *Let  $\mathfrak{J}$  be a finite AVQC with  $\overline{C}_{\text{random}}(\mathfrak{J}, 0) > 0$  and  $\lambda, \varepsilon > 0$ ,  $0 < R < \overline{C}_{\text{random}}(\mathfrak{J}, 0)$ . There exist an  $L(\mathfrak{J}, \lambda, R, \varepsilon) \in \mathbb{N}$  and  $K(\mathfrak{J}, \lambda, R, \varepsilon) \in \mathbb{N}$  such that for all  $l \geq L(\mathfrak{J}, \lambda, R, \varepsilon)$  there are  $K(\mathfrak{J}, \lambda, R, \varepsilon)$  deterministic codes for the AVQC  $\mathfrak{J}$  such that*

$$\frac{1}{K(\mathfrak{J}, \lambda, R, \varepsilon)} \sum_{j=1}^{K(\mathfrak{J}, \lambda, R, \varepsilon)} \frac{1}{M_l} \sum_{i=1}^{M_l} \text{tr}(\mathcal{N}_{s^l}(\rho_{i,j}) D_{i,j}^l) \geq 1 - \lambda \quad \forall s^l \in \mathbf{S}^l, \quad \frac{1}{l} \log M_l \geq R - \varepsilon. \quad (24)$$

It holds, with the abbreviation  $E := \overline{E}_{\text{m}}(\text{conv}(\mathfrak{J}), R)$ ,

$$L(\mathfrak{J}, \lambda, R) = \min\{L : L(1 - \frac{2 \cdot |\mathbf{S}|}{E - \varepsilon} \cdot \frac{1}{L} \log(L) \geq \frac{2}{E - \varepsilon} \log(\frac{1}{\lambda} \cdot \frac{4}{E - \varepsilon})\}, \quad (25)$$

and the quantity  $K(\mathfrak{J}, \lambda, R)$  is given by

$$K(\mathfrak{J}, \lambda, R, \varepsilon) = \frac{1}{\lambda} \cdot \frac{8 \cdot \log |\mathbf{S}|}{E - \varepsilon}. \quad (26)$$

The same statement holds with average error criterion replaced by entanglement fidelity and  $\overline{E}_{\text{m}}$  by  $E_{\text{e}}$ .

**Remark 7.** It is clear that above statement is especially interesting for the message transmission capacity of an AVQC, and there only in the case when the deterministic capacity vanishes but the randomness assisted one does not.

As a very rough approximation, one may use the scaling law  $L(\mathfrak{J}, \lambda, R) \approx \frac{2}{\overline{E}_{\text{m}}(\text{conv}(\mathfrak{J}), R) - \varepsilon} \log(\frac{1}{\lambda} \cdot \frac{4}{\overline{E}_{\text{m}}(\text{conv}(\mathfrak{J}), R) - \varepsilon})$ . It is clear that both  $L$  and  $K$  from above theorem are sub-optimal, even with the techniques used in this paper. However, their scaling with  $\lambda$  does not depend on the choice of constants in our proof. For fixed  $\mathfrak{J}$  and rate  $R$ , this means that the block-length needed to achieve a certain error  $\lambda$  roughly scales as  $\log(1/\lambda)$ , and the randomness as  $1/\lambda$ .

**Theorem 7.** *Let  $\mathfrak{I}$  be a finite AVQC and  $\lambda \in [0, 1]$ . Then both*

$$\mathcal{A}_{\text{det}}(\mathfrak{I}, \lambda) = \mathcal{A}_{\text{s,det}}(\mathfrak{I}, \lambda) \quad \text{and} \quad \mathcal{A}_{\text{random}}(\mathfrak{I}, \lambda) = \mathcal{A}_{\text{s,random}}(\mathfrak{I}, \lambda). \quad (27)$$

**Remark 8.** *It should, of course, be noted that we expect this picture to change once finite block-lengths are brought into the game. We leave this for future work.*

## V Proofs

We now give the proofs of our results, in the same order they appeared in Section IV.

### V.1 Using randomness is advantageous for message transmission over AVQCs

This subsection is devoted to the proof of Theorem 2.

*Proof.* We will first construct an AVQC having the special structure introduced in Theorem 2, then prove that every AVQC having that special form is symmetrizable, and finally estimate the random capacity for the one special choice we made:

1) A particular example is given by  $\mathbf{X} = \mathbf{S} = \{0, 1\}$  and with  $\mathcal{K} = \mathbb{C}^3$ . Take an arbitrary POVM  $\{M_x\}_{x \in \mathbf{X}}$  and three different states  $\sigma_1, \sigma_2, \sigma_3$ . This will be enough to prove that  $\overline{\mathcal{C}}_{\text{det}}(\mathfrak{I}, 0) = 0$ . In order to get an easy estimate on the random capacity, we have to make a more explicit choice for the  $\sigma_i$ . The choice  $\sigma_1 = |e_1\rangle\langle e_1|$ ,  $\sigma_2 = |e_2\rangle\langle e_2|$ ,  $\sigma_3 = |e_3\rangle\langle e_3|$ ,  $M_i = |e_i\rangle\langle e_i|$ ,  $i = 1, 2$  will do. Set

$$\rho_{1,1} = \sigma_1, \quad \rho_{1,2} = \sigma_3, \quad \rho_{2,1} = \sigma_3, \quad \rho_{2,2} = \sigma_2. \quad (28)$$

Then obviously the choice  $p_1(i) = \delta(1, i)$ ,  $p_2(i) = \delta(2, i)$  (with  $\delta(\cdot, \cdot)$  being the usual Kronecker-delta function) fulfills equation (20).

2)  $\mathfrak{I}$  is 1-symmetrizable: Given the states  $\{\nu_x\}_{x \in \mathbf{X}}$ , we choose the distributions  $\{q_x\}_{x \in \mathbf{X}}$  defined by

$$q_x(s) := \sum_{\tilde{x} \in \mathbf{X}} p_{\tilde{x}}(s) \text{tr}\{M_{\tilde{x}} \nu_x\}, \quad (29)$$

then symmetrizability holds by the following calculation:

$$\sum_{s \in \mathbf{S}} q_{x'}(s) W_s(\nu_x) = \sum_{s \in \mathbf{S}} q_{x'}(s) \sum_{\tilde{x} \in \mathbf{X}} \text{tr}\{\nu_x M_{\tilde{x}}\} \rho_{s, \tilde{x}} \quad (30)$$

$$= \sum_{\tilde{x} \in \mathbf{X}} \text{tr}\{\nu_x M_{\tilde{x}}\} \sum_{s \in \mathbf{S}} q_{x'}(s) \rho_{s, \tilde{x}} \quad (31)$$

$$= \sum_{\tilde{x} \in \mathbf{X}} \text{tr}\{\nu_x M_{\tilde{x}}\} \sum_{s \in \mathbf{S}} \sum_{\hat{x} \in \mathbf{X}} p_{\hat{x}}(s) \text{tr}\{M_{\hat{x}} \nu_{x'}\} \rho_{s, \tilde{x}} \quad (32)$$

$$= \sum_{\tilde{x} \in \mathbf{X}} \text{tr}\{\nu_x M_{\tilde{x}}\} \sum_{\hat{x} \in \mathbf{X}} \text{tr}\{M_{\hat{x}} \nu_{x'}\} \sum_{s \in \mathbf{S}} p_{\hat{x}}(s) \rho_{s, \tilde{x}} \quad (33)$$

$$= \sum_{\tilde{x} \in \mathbf{X}} \text{tr}\{\nu_x M_{\tilde{x}}\} \sum_{\hat{x} \in \mathbf{X}} \text{tr}\{M_{\hat{x}} \nu_{x'}\} \sum_{s \in \mathbf{S}} p_{\hat{x}}(s) \rho_{s, \tilde{x}} \quad (34)$$

$$= \sum_{s \in \mathbf{S}} \left( \sum_{\tilde{x} \in \mathbf{X}} \text{tr}\{\nu_x M_{\tilde{x}}\} p_{\tilde{x}}(s) \right) \sum_{\hat{x} \in \mathbf{X}} \text{tr}\{M_{\hat{x}} \nu_{x'}\} \rho_{s, \hat{x}} \quad (35)$$

$$= \sum_{s \in \mathbf{S}} q_x(s) \sum_{\hat{x} \in \mathbf{X}} \text{tr}\{M_{\hat{x}} \nu_{x'}\} \rho_{s, \hat{x}} \quad (36)$$

$$= \sum_{s \in \mathbf{S}} q_x(s) W_s(\nu_{x'}). \quad (37)$$

3) We now want to get a nonzero lower bound on the random capacity of  $\overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0)$ . To do so, we employ the results [5] of Ahlswede and Blinovsky:

Fix the allowed input states  $\mathcal{P}(i)$  in the encoding of messages into  $\mathfrak{J}$  to arbitrary tensor products of a set  $\{\rho_1, \dots, \rho_{\dim(\mathcal{H})}\}$  of pure states. It is clear that such a strategy is hopelessly sub-optimal. But by restriction to this model, we get from Theorem 1 in [5] the lower bound

$$\overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0) \geq \max_{p \in \mathfrak{P}([\dim(\mathcal{H})])} \min_{\mathcal{N} \in \text{conv}(\mathfrak{J})} S\left(\sum_{k=1}^K p(k) \mathcal{N}(\rho_k)\right) - \sum_{k=1}^K p(k) S(\mathcal{N}(\rho_k)). \quad (38)$$

But the function  $(p, \mathcal{N}) \rightarrow S(\sum_{k=1}^K p(k) \mathcal{N}(\rho_k)) - \sum_{k=1}^K p(k) S(\mathcal{N}(\rho_k))$  is concave in  $p$  and convex in  $\mathcal{N}$  and both optimization procedures in (38) are over convex compact sets, so this translates to

$$\overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0) \geq \min_{\mathcal{N} \in \text{conv}(\mathfrak{J})} \max_{p \in \mathfrak{P}([K])} S\left(\sum_{k=1}^K p(k) \mathcal{N}(\rho_k)\right) - \sum_{k=1}^K p(k) S(\mathcal{N}(\rho_k)), \quad (39)$$

by the minimax-Theorem. Therefore,

$$\overline{\mathcal{C}}_{\text{random}}(\mathfrak{W}, 0) \geq \min_{\mathcal{N} \in \text{conv}(\mathfrak{J})} \max_{\{q_i, \sigma_i\}} \chi(\{q_i, \sigma_i\}_i, \mathcal{N}) \quad (40)$$

$$\geq \min_{\mathcal{N} \in \text{conv}(\{\mathcal{N}_1, \mathcal{N}_2\})} \chi\left(\left\{\frac{1}{2}, |e_i\rangle\langle e_i|_{i=1}^2\right\}, \mathcal{N}\right) \quad (41)$$

$$= \min_{t \in [0,1]} H([t/2, (1-t)/2, 1/2, 0]) - t \cdot h(1/2) - (1-t) \cdot h(1/2) \quad (42)$$

$$= 1 - h(t)/2 \quad (43)$$

$$\geq 1/2 \quad (44)$$

$$> 0. \quad (45)$$

□

We use the opportunity to state the following additional Lemma:

**Lemma 2** (Message Transmission Capacity of an AVQC). *Let  $\mathfrak{J}$  be a finite AVQC. Then*

$$\overline{\mathcal{C}}_{\text{random}}(\mathfrak{J}, 0) = \lim_{l \rightarrow \infty} \frac{1}{l} \min_{\mathcal{N} \in \text{conv}(\mathfrak{J})} \chi(\mathcal{N}^{\otimes l}). \quad (46)$$

*Proof.* The converse statement “LHS  $\leq$  RHS” follows from observing first that each code for the AVQC  $\mathfrak{J}$  is also a code for the compound quantum channel  $\mathfrak{J}$ , with the same average error.

The inequality “LHS  $\geq$  RHS” follows from using any of the existing proofs for the converse theorem for the memoryless stationary quantum channel in combination with the minimax theorem.

The direct part is a consequence of [5], together with the usual blocking strategies. For more information on how these steps are carried out in detail, we refer to [7], where this is done for entanglement transmission. □

## V.2 Investigation of Continuity

In this subsection, we prove Theorem 3, Theorem 5, Theorem 4, its corollaries and, finally, Lemma 1.

Let us get started. It is important to note here, that a crucial ingredient to the proof of Theorem 2 is that  $D_\eta$  is *not* entanglement breaking for  $\eta \neq 1$ , while  $\mathfrak{J}$  from Theorem 2 is. At the same time, the  $D_\eta$  have a very simple structure. This is what enables us to prove non-symmetrizability for all  $\lambda, \eta \in [0, 1]$ :

*Proof of Theorem 3.* By Theorem 2 we know that  $\mathfrak{J}_1^\eta = \mathfrak{J}$  is symmetrizable for all  $\eta \in [0, 1]$ . It is, as an additional fact, clear that  $\mathfrak{J}_\lambda^1$  is symmetrizable for all  $\lambda \in [0, 1]$ .

We are going to show that  $\mathfrak{J}_\lambda^\eta$  is non-symmetrizable for all  $\lambda, \eta \in [0, 1]$ :

Assume that, to a given pair  $\rho, \sigma \in \mathcal{S}(\mathbb{C}^2)$  of input states there are  $p, q \in \mathcal{P}(\mathbf{S})$  such that

$$\sum_{s \in \mathbf{S}} p(s) \hat{\mathcal{N}}_{s, \eta, \lambda}(\rho) = \sum_{s \in \mathbf{S}} q(s) \hat{\mathcal{N}}_{s, \eta, \lambda}(\sigma). \quad (47)$$

Rearranging terms, we get that this is equivalent to

$$\lambda \sum_{s \in \mathbf{S}} [q(s) \mathcal{N}_s(\sigma) - p(s) \mathcal{N}_s(\rho)] - (1 - \lambda) \eta \pi = (1 - \lambda)(1 - \eta)(\rho - \sigma). \quad (48)$$

Epecially, the latter equality would have to hold for the entries of the respective matrices in the standard orthonormal basis  $\{e_i\}_{i=1}^3$  we may choose, so that a look at the off-diagonal entries reveals that (since the channels  $\mathcal{N}_s$  are all entanglement-breaking), the equality

$$0 = (1 - \lambda)(1 - \eta) \langle e_1, (\rho - \sigma) e_2 \rangle \quad (49)$$

would have to hold. Since both  $\lambda \neq 1$  and  $\eta \neq 1$  we see that this would require  $\langle e_1, (\rho - \sigma) e_2 \rangle = 0$ , and that is clearly not valid for all choices  $\rho, \sigma$  of input states. Thus  $\mathfrak{J}_\lambda^\eta$  is non-symmetrizable for  $\lambda, \eta \in [0, 1]$ .  $\square$

We now turn to the proof of Theorem 4 and its corollaries, after which we give a proof of Theorem 5. Then, we prove Lemma 1. A preliminary definition is needed:

**Definition 10.** To any given  $l \in \mathbb{N}$ , define a nonnegative function  $F_l$  on the set of all finite subsets of  $\mathcal{C}(\mathcal{H}^{\otimes l}, \mathcal{K}^{\otimes l})$  through

$$\{\mathcal{N}_i\}_{i \in \mathcal{I}} \mapsto \max_{\rho, \sigma \in \mathcal{S}(\mathcal{H}^{\otimes l})} \min_{q, p \in \mathcal{P}(\mathcal{I})} \left\| \sum_{i \in \mathcal{I}} [p(i) \mathcal{N}_i(\rho) - q(i) \mathcal{N}_i(\sigma)] \right\|_1. \quad (50)$$

**Remark 9.** This function is well-defined, by compactness of  $\mathcal{S}(\mathcal{H}^{\otimes l})$  and  $\mathcal{P}(\mathcal{I})$ .

**Remark 10.** For any finite AVQC  $\mathfrak{J}$  the statements " $\mathfrak{J}$  is  $l$ -symmetrizable" and " $F_l(\mathfrak{J}) = 0$ " are equivalent.

*Proof of Theorem 4.* Let  $\mathfrak{J} = \{\mathcal{N}_s\}_{s \in \mathbf{S}}$ . Let  $q, p \in \mathcal{P}(\mathbf{S})$  be arbitrary and  $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ . Let  $\mathfrak{J}'$  be a finite AVQC satisfying

$$D_\diamond(\mathfrak{J}, \mathfrak{J}') \leq \eta \quad (51)$$

for some  $\eta > 0$  that will be specified later. We write  $\mathfrak{J}' = \{\mathcal{N}'_x\}_{x \in \mathbf{X}}$ , and by using the same channel twice if necessary we may assume that  $\mathbf{X} = \mathbf{S}$  holds. The inequality (51) implies that there are functions  $f, g : \mathbf{S} \rightarrow \mathbf{S}$  such that for every  $s \in \mathbf{S}$  we have

$$\|\mathcal{N}'_s - \mathcal{N}_{f(s)}\|_\diamond \leq \eta, \quad \|\mathcal{N}'_{g(s)} - \mathcal{N}_s\|_\diamond \leq \eta, \quad \forall s \in \mathbf{S}. \quad (52)$$

It then follows from convexity of  $\|\cdot\|_\diamond$  that

$$\left\| \sum_{s \in \mathbf{S}} p(s) [\mathcal{N}_s - \mathcal{N}'_{g(s)}] \right\|_\diamond \leq \eta, \quad (53)$$

$$\left\| \sum_{s \in \mathbf{S}} q(s) [\mathcal{N}_s - \mathcal{N}'_{f(s)}] \right\|_\diamond \leq \eta. \quad (54)$$

Define  $\tilde{q}, \tilde{p} \in \mathcal{P}(\mathbf{S})$  by

$$\tilde{q}(s) := \sum_{s' : g(s') = s} q(s'), \quad \tilde{p}(s) := \sum_{s' : f(s') = s} p(s'). \quad (55)$$

It then holds that

$$\sum_{s \in \mathbf{S}} q(s) \mathcal{N}_{g(s)} = \sum_{s \in \mathbf{S}} \tilde{q}(s) \mathcal{N}_s, \quad \sum_{s \in \mathbf{S}} p(s) \mathcal{N}_{f(s)} = \sum_{s \in \mathbf{S}} \tilde{p}(s) \mathcal{N}_s. \quad (56)$$

Thus

$$\left\| \sum_{s \in \mathbf{S}} [\tilde{q}(s) \mathcal{N}_s(\sigma) - \tilde{p}(s) \mathcal{N}_s(\rho)] \right\|_1 = \left\| \sum_{s \in \mathbf{S}} [\tilde{q}(s) \mathcal{N}_s(\sigma) - q(s) \mathcal{N}'_s(\sigma) + q(s) \mathcal{N}'_s(\sigma) + \right. \quad (57)$$

$$\left. + p(s) \mathcal{N}'_s(\rho) - p(s) \mathcal{N}'_s(\rho) - \tilde{p}(s) \mathcal{N}_s(\rho)] \right\|_1 \quad (58)$$

$$\leq \left\| \sum_{s \in \mathbf{S}} [\tilde{q}(s) \mathcal{N}_s(\sigma) - q(s) \mathcal{N}'_s(\sigma)] \right\|_1 + \left\| \sum_{s \in \mathbf{S}} [p(s) \mathcal{N}'_s(\rho) - \right. \quad (59)$$

$$\left. - \tilde{p}(s) \mathcal{N}_s(\rho)] \right\|_1 + \left\| \sum_{s \in \mathbf{S}} [q(s) \mathcal{N}'_s(\sigma) - p(s) \mathcal{N}'_s(\rho)] \right\|_1 \quad (60)$$

$$\leq 2\eta + \left\| \sum_{s \in \mathbf{S}} [q(s) \mathcal{N}'_s(\sigma) - p(s) \mathcal{N}'_s(\rho)] \right\|_1. \quad (61)$$

This holds for all  $\rho, \sigma, p, q$ , and minimizing the left hand side we especially get that

$$\min_{q', p' \in \mathcal{P}(\mathbf{S})} \left\| \sum_{s \in \mathbf{S}} [q'(s) \mathcal{N}_s(\sigma) - p'(s) \mathcal{N}_s(\rho)] \right\|_1 \leq 2\eta + \left\| \sum_{s \in \mathbf{S}} [q(s) \mathcal{N}'_s(\sigma) - p(s) \mathcal{N}'_s(\rho)] \right\|_1. \quad (62)$$

This implies

$$\min_{q, p \in \mathcal{P}(\mathbf{S})} \left\| \sum_{s \in \mathbf{S}} [q(s) \mathcal{N}_s(\sigma) - p(s) \mathcal{N}_s(\rho)] \right\|_1 \leq 2\eta + \min_{q, p \in \mathcal{P}(\mathbf{S})} \left\| \sum_{s \in \mathbf{S}} [q(s) \mathcal{N}'_s(\sigma) - p(s) \mathcal{N}'_s(\rho)] \right\|_1, \quad (63)$$

and from there it clearly follows that  $F_1(\mathcal{J}) \leq 2\eta + F_1(\mathcal{J}')$ . Now choose  $\eta = F_1(\mathcal{J})/4$ , then we get  $\blacksquare$   $F_1(\mathcal{J}) \leq F_1(\mathcal{J})/2 + F_1(\mathcal{J}')$ . But that directly implies

$$F_1(\mathcal{J})/2 \leq F_1(\mathcal{J}'), \quad (64)$$

and this in turn implies  $\overline{C}_{\det}(\mathcal{J}', 0) > 0$ .  $\square$

*Proof of Corollary 1.* It is clear that there is an  $L \in \mathbb{N}$  such that for all  $l \geq L$  we have  $\overline{C}_{\det}(\mathcal{J}_l, 0) > 0$ . But if  $\overline{C}_{\det}(\mathcal{J}_l, 0) > 0$ , then  $\overline{C}_{\det}(\mathcal{J}_l, 0) = \overline{C}_{\text{random}}(\mathcal{J}_l, 0)$  and the latter quantity is continuous by the formula in Lemma 2 and the results [27] of Leung and Smith (for their application to quantum capacity formulae of AVQCs, [7] is a good reference), so (reading equalities from left to right)

$$\lim_{l \rightarrow \infty} \overline{C}_{\det}(\mathcal{J}_l, 0) = \lim_{l \rightarrow \infty} \overline{C}_{\text{random}}(\mathcal{J}_l, 0) = \overline{C}_{\text{random}}(\mathcal{J}, 0). \quad (65)$$

$\square$

*Proof of Corollary 2.* The l.h.s. equality has been proven in ([7], Theorem 1), and the r.h.s. inequality is proven in the same way as Corollary 1: For all  $l \geq L$ , we have that  $\mathcal{A}_{\det}(\mathcal{J}_l) = \mathcal{A}_{\text{random}}(\mathcal{J}_l)$ , and the latter quantity is continuous with respect to  $D_\diamond$  (this was implicitly proven in [7], see equations (98) and (99)), proving the corollary.  $\square$

In the following lines, we prove Theorem 5.

*Proof of Theorem 5.* Let  $\mathcal{J}$  ( $|\mathcal{J}| < \infty!$ ) be a discontinuity point of  $\overline{C}_{\det}(\cdot, 0)$ . By Corollary 1 we know that  $\overline{C}_{\det}(\mathcal{J}, 0) = 0$  via contradiction. This implies that there exists a  $\delta > 0$  such that  $\overline{C}_{\det}(\mathcal{J}', 0) \geq \delta$  for all  $\mathcal{J}'$  satisfying  $0 < D_\diamond(\mathcal{J}', \mathcal{J}) \leq \delta$ .

We still have to show that  $\overline{C}_{\text{random}}(\mathcal{J}, 0) > 0$  follows, so assume the contrary:  $\overline{C}_{\text{random}}(\mathcal{J}, 0) = 0$ . Then, there would certainly exist a finite AVQC  $\mathcal{J}'$  satisfying  $D_\diamond(\mathcal{J}', \mathcal{J}) \leq \delta$  and  $\overline{C}_{\text{random}}(\mathcal{J}', 0) < \delta$ . But then

$$\delta \leq \overline{C}_{\det}(\mathcal{J}', 0) \leq \overline{C}_{\text{random}}(\mathcal{J}', 0) < \delta \quad (66)$$

would have to hold, a clear contradiction. Thus,  $\overline{C}_{\text{random}}(\mathfrak{J}, 0) > 0$ .

On the other hand, let  $\overline{C}_{\text{det}}(\mathfrak{J}, 0) = 0$ ,  $\overline{C}_{\text{random}}(\mathfrak{J}, 0) > 0$  and assume that for all  $\delta > 0$  there exists a finite  $\mathfrak{J}_\delta$  satisfying  $0 < D_\diamond(\mathfrak{J}_\delta, \mathfrak{J}) \leq \delta$  such that  $\overline{C}_{\text{det}}(\mathfrak{J}_\delta, 0) > 0$ . Then by continuity of  $\overline{C}_{\text{random}}(\cdot, 0)$  we have for some  $\varepsilon > 0$  that  $\overline{C}_{\text{random}}(\mathfrak{J}_\delta, 0) > \varepsilon$  for all those  $\mathfrak{J}_\delta$  for which  $\delta$  is small enough ( $0 < \delta < \delta_0$ , say). In addition, it is clear that for  $0 < \delta < \delta_0$  we also have  $\overline{C}_{\text{det}}(\mathfrak{J}_\delta, 0) = \overline{C}_{\text{random}}(\mathfrak{J}_\delta, 0)$ , whence  $\overline{C}_{\text{det}}(\mathfrak{J}_\delta, 0) > \varepsilon$ . This proves that  $\overline{C}_{\text{det}}(\cdot, 0)$  is discontinuous in the point  $\mathfrak{J}$ .  $\square$

*Proof of Lemma 1.* Clearly, by assumption we have that  $\mathfrak{J}$  is  $l$ -symmetrizable. Fix  $\sigma \in \mathcal{S}(\mathcal{H}^{\otimes l})$  and define  $\mathcal{N}_\sigma := \sum_{s^l \in \mathbf{S}^l} p[\sigma](s^l) \mathcal{N}_{s^l}$ . Then it holds, for all  $\rho \in \mathcal{S}(\mathcal{H}^{\otimes l})$ ,

$$\mathcal{N}_\sigma(\rho) = \sum_{s^l \in \mathbf{S}^l} p[\sigma](s^l) \mathcal{N}_{s^l}(\rho) = \sum_{s^l \in \mathbf{S}^l} q[\rho](s^l) \mathcal{N}_{s^l}(\sigma). \quad (67)$$

Especially for every two states  $\rho, \rho'$  and  $\lambda \in (0, 1)$  we get

$$\sum_{s^l \in \mathbf{S}^l} q[\lambda\rho + (1-\lambda)\rho'](s^l) \mathcal{N}_{s^l}(\sigma) = \sum_{s^l \in \mathbf{S}^l} (\lambda q[\rho](s^l) + (1-\lambda)q[\rho'](s^l)) \mathcal{N}_{s^l}(\sigma). \quad (68)$$

The operators  $\mathcal{N}_{s^l}(\sigma)$  may be linearly dependent, so that no general conclusion can be drawn yet, we have to go one step further: The set  $\text{conv}(\mathcal{N}_{s^l}(\sigma))$  has extremal points  $\{\mathcal{N}_t\}_{t \in T}$  for some subset  $T \subset \mathbf{S}^l$  and we may choose probability distributions  $\{r(\cdot|s^l)\}_{s^l \in \mathbf{S}^l} \subset \mathfrak{P}(T)$  such that we get

$$\forall s^l \in \mathbf{S}^l : \quad \mathcal{N}_{s^l}(\sigma) = \sum_{t \in T} r(t|s^l) \mathcal{N}_t(\sigma). \quad (69)$$

Defining  $\tilde{q} : \mathcal{S}(\mathcal{H}^{\otimes l}) \rightarrow \mathfrak{P}(T)$  pointwise through  $\tilde{q}[\rho](t) := \sum_{s^l \in \mathbf{S}^l} q[\rho](s^l) r(t|s^l)$  this transforms the statement (68) into

$$\sum_{t \in T} \tilde{q}[\lambda\rho + (1-\lambda)\rho'](t) \mathcal{N}_t(\sigma) = \sum_{t \in T} (\lambda \tilde{q}[\rho](t) + (1-\lambda)\tilde{q}[\rho'](t)) \mathcal{N}_t(\sigma). \quad (70)$$

Since convex combinations of extremal points are unique, this proves that for all  $\rho, \rho' \in \mathcal{S}(\mathcal{H}^{\otimes l})$  and  $\lambda \in (0, 1)$  we have

$$\tilde{q}[\lambda\rho + (1-\lambda)\rho'] = \lambda \tilde{q}[\rho] + (1-\lambda)\tilde{q}[\rho']. \quad (71)$$

In a next step we can now linearly extend the functional  $\tilde{q}$  to all of  $\mathcal{B}(\mathcal{H}^{\otimes l})$ : the (complex) linear span of  $\mathcal{S}(\mathcal{H}^{\otimes l})$  obviously contains the set of all nonnegative operators, whose linear span clearly contains the self-adjoint operators, and their linear span finally is  $\mathcal{B}(\mathcal{H}^{\otimes l})$ .

We have thus (since  $T \subset \mathbf{S}^l$ ) constructed a linear functional  $\tilde{q}$  such that  $\tilde{q}(\mathcal{S}(\mathcal{H}^{\otimes l})) \subset \mathfrak{P}(\mathbf{S}^l)$  and

$$\forall \rho \in \mathcal{S}(\mathcal{H}^{\otimes l}) : \quad \mathcal{N}_\sigma(\rho) = \sum_{s^l \in \mathbf{S}^l} \tilde{q}[\rho](s^l) \mathcal{N}_{s^l}(\sigma). \quad (72)$$

Since each  $\tilde{q}[\cdot](t)$  is a linear functional it may be represented as  $\tilde{q}[\cdot](t) = \text{tr}\{A_t \cdot\}$  for some operator  $A_t$ , and since it is both positivity preserving and trace preserving the operators  $A_t$  can be chosen to be nonnegative and must satisfy the condition  $\sum_{t \in T} A_t = \mathbb{1}_{\mathcal{H}^{\otimes l}}$ , whence the form a POVM and it holds

$$\forall \rho \in \mathcal{S}(\mathcal{H}^{\otimes l}) : \quad \mathcal{N}_\sigma(\rho) = \sum_{s^l \in \mathbf{S}^l} \text{tr}\{A_{s^l} \rho\} \mathcal{N}_{s^l}(\sigma). \quad (73)$$

This is equivalent to the statement that  $\mathcal{N}_\sigma \in \text{conv}(\{\mathcal{N}_{s^l}\})$  is entanglement breaking by the results of [23], which implies that  $\mathcal{A}_{\text{random}}(\{\mathcal{N}_{s^l}\}_{s^l \in \mathbf{S}^l}) = 0$  by the results of [7], which is equivalent to  $\mathcal{A}_{\text{random}}(\mathfrak{J}, 0) = 0$  by the very definition of the channel model.  $\square$

### V.3 Finite amounts of randomness, finite errors and blocklengths

In this subsection, we give the proofs of the results concerning finite resources. Namely, we are interested here in the scaling laws that connect finite errors  $\lambda$  and the amount of randomness needed to achieve such an error. At the same time, we are able to give estimates on the block-length that is needed to achieve a certain error.

In addition, we present the following Lemma which is an immediate drop-off from the results in [7]:

**Lemma 3** (A partial strong converse for message transmission over finite AVQCs). *For every finite AVQC  $\mathfrak{J}$  which is symmetrizable and for every  $\lambda \in [0, 1/2)$  we have  $\overline{C}(\mathfrak{J}, \lambda) = \overline{C}(\mathfrak{J}, 0)$ .*

*Proof.* This follows directly from the proof of Theorem 40 in [7], with a slightly better estimate in between equations (202) and (203).  $\square$

We now turn to the original goal of this subsection, the proof of Theorem 6:

*Proof of Theorem 6.* Let  $l \in \mathbb{N}$ ,  $\varepsilon > 0$  and  $\delta := \overline{E}_m(\text{conv}(\mathfrak{J}), \overline{C}_{\text{compound}}(\text{conv}(\mathfrak{J}))) - \varepsilon$ . If only  $l$  is large enough, then by definition of  $E$  there is a code  $\{\sigma_i, P_i\}_{i=1}^{M_l}$  with  $M_l = \lfloor 2^{l(\overline{C}_{\text{compound}}(\text{conv}(\mathfrak{J})) - \varepsilon)} \rfloor$  such that

$$\frac{1}{M_l} \sum_{i=1}^{M_l} \text{tr}\{P_i \mathcal{N}^{\otimes n}(\sigma_i)\} \geq 1 - 2^{-l\delta} \quad \forall \mathcal{N} \in \text{conv}(\mathfrak{J}) \quad (74)$$

and that  $\delta$  is the largest possible such value over all choices of codes at the same rate and same blocklength. By application of the robustification technique (find the original in [3], Theorem 6, an improved version in [4] or read Theorem 28 in [7]) we get a random code  $\mu_l$  at the same rate such that

$$e(\mu_l, \mathcal{N}) := \min_{s^l \in \mathbf{S}^l} \int \frac{1}{M_l} \sum_{i=1}^{M_l} \text{tr}(\mathcal{N}_{s^l}(\rho_i) D_i^l) d\mu_l((\rho_i, D_i^l)_{i=1}^{M_l}) \geq 1 - (l+1)^{|\mathbf{S}|} 2^{-l\delta}. \quad (75)$$

We will now use the abbreviation  $\varepsilon_l := (l+1)^{|\mathbf{S}|} 2^{-l\delta}$ . For a fixed  $K \in \mathbb{N}$ , consider  $K$  independent random variables  $\Lambda_i$  with values in  $((\mathcal{S}(\mathcal{H}^{\otimes l})^{M_l}) \times \mathcal{M}_{M_l}(\mathcal{H}^{\otimes l}))$  which are distributed according to  $\mu_l$ .

Define, for each  $s^l \in \mathbf{S}^l$ , the function  $p_{s^l} : ((\mathcal{S}(\mathcal{H}^{\otimes l})^{M_l}) \times \mathcal{M}_{M_l}(\mathcal{H}^{\otimes l})) \rightarrow [0, 1]$ ,

$$(\rho_1, \dots, \rho_{M_l}, D_1^l, \dots, D_{M_l}^l) \mapsto \frac{1}{M_l} \sum_{i=1}^{M_l} \text{tr}(\mathcal{N}_{s^l}(\rho_i) D_i^l).$$

We get, by application of Markov's inequality, for every  $s^l \in \mathbf{S}^l$ , and every  $r \geq 0$ :

$$\mathbb{P}(1 - \frac{1}{K} \sum_{j=1}^K p_{s^l}(\Lambda_j) \geq \lambda) = \mathbb{P}(2^{rK-r \sum_{j=1}^K p_{s^l}(\Lambda_j)} \geq 2^{rK\lambda}) \quad (76)$$

$$\leq 2^{-rK\lambda} \mathbb{E}(2^{(rK-r \sum_{j=1}^K p_{s^l}(\Lambda_j))}). \quad (77)$$

The  $\Lambda_i$  are independent and it holds  $2^{rt} \leq 1 + t2^r$  for every  $t \in [0, 1]$  and  $r \geq 0$  as well as  $\log(1+x) \leq 2x$  for  $x \geq 0$  and so we get

$$\mathbb{P}(1 - \frac{1}{K} \sum_{j=1}^K p_{s^l}(\Lambda_j) \geq \lambda) \leq 2^{-rK\lambda} \mathbb{E}(2^{rK-r \sum_{j=1}^K p_{s^l}(\Lambda_j)}) \quad (78)$$

$$= 2^{-rK\lambda} \mathbb{E}(2^{r \sum_{j=1}^K (1-p_{s^l}(\Lambda_j))}) \quad (79)$$

$$= 2^{-rK\lambda} \mathbb{E}(2^{(r-r p_{s^l}(\Lambda_1))})^K \quad (80)$$

$$\leq 2^{-rK\lambda} \mathbb{E}(1 + (1 - p_{s^l}(\Lambda_1) 2^r))^K \quad (81)$$

$$\leq 2^{-rK\lambda} \mathbb{E}(1 + \varepsilon_l 2^r)^K \quad (82)$$

$$\leq 2^{-rK\lambda} 2^{K\varepsilon_l 2^r} \quad (83)$$

$$= 2^{-K(r\lambda - \varepsilon_l 2^r)}. \quad (84)$$



Therefore, and with the choice  $r = l\delta/2$ , we get by application of the union bound that

$$\mathbb{P}\left(\frac{1}{K} \sum_{j=1}^K p_{s^l}(\Lambda_j) \geq 1 - \lambda \ \forall s^l \in \mathbf{S}^l\right) \geq 1 - |\mathbf{S}|^l 2^{-K(l\delta\lambda/2 - (l+1)|\mathbf{S}|2^{-l\delta/2})} \quad (85)$$

$$= 1 - 2^{-l(K(\delta\lambda/2 - \frac{(l+1)|\mathbf{S}|}{l}2^{-l\delta/2}) - \log(|\mathbf{S}|))}. \quad (86)$$

But this shows that, for  $K > \log(|\mathbf{S}|)/(\frac{\delta\lambda}{2} - \frac{(l+1)|\mathbf{S}|}{l}2^{-l\delta/2})$ , the above probability is larger than zero, so there exists a realization  $\Lambda_1, \dots, \Lambda_{l^n}$  such that

$$\min_{s^l \in \mathbf{S}^l} \frac{1}{K} \sum_{i=1}^K \frac{1}{M_l} \text{tr}(\mathcal{N}_{s^l}(\rho_i) D_i^l) \geq 1 - \lambda. \quad (87)$$

The important conclusion we draw from this result is that, for  $l$  large enough ( $l \geq L(\delta, \lambda, |\mathbf{S}|)$  for some  $L(\delta, \lambda, |\mathbf{S}|) \in \mathbb{N}$ ),  $\frac{(l+1)|\mathbf{S}|}{l}2^{-l\delta/2} < \delta\lambda/4$ , and then for  $K = 8 \log(|\mathbf{S}|)/\delta\lambda$  we are the guaranteed existence of random codes that only use a finite amount of randomness. This proves the theorem.  $\square$

*Proof of Theorem 7.* At last, let us consider the differences between entanglement- and strong subspace transmission. Let  $l \in \mathbb{N}$  be arbitrary. It has been shown in [7], their Lemma 21, that every subspace  $\mathcal{F}_l$  for entanglement transmission (with shared randomness or without) over an (even non-finite) AVQC  $\mathfrak{J}$  with error  $\lambda_l \in [0, 1]$  contains a subspace  $\hat{\mathcal{F}}_l \subset \mathcal{F}_l$  for strong subspace transmission over  $\mathfrak{J}$  with error  $\hat{\lambda}_l \leq \lambda_l + c/\sqrt{k_l - 1} + \varepsilon_l$ , where  $k_l = \dim \mathcal{F}_l$  and  $\dim \hat{\mathcal{F}}_l = \lfloor \frac{\varepsilon_l^2}{256 \log(32/\varepsilon_l)} k_l \rfloor$ , while  $\varepsilon_l$  can be suitably chosen.

These results were then used in [7] to conclude that, with a specific choice of  $\varepsilon_l$  and with  $R := \liminf_{l \rightarrow \infty} \frac{1}{l} \log k_l$ , one can obtain an error  $\hat{\lambda}_l \leq \lambda_l + c/\sqrt{k_l - 1} + 2^{-l \cdot R/4}$  and such that  $R = \liminf_{l \rightarrow \infty} \frac{1}{l} \log \hat{k}_l$ .

The implication for the  $\lambda$ -capacities is that the following estimates hold:

$$\mathcal{A}_{\text{s,det}}(\mathfrak{J}, \lambda) \geq \mathcal{A}_{\text{det}}(\mathfrak{J}, \lambda) \quad \text{and} \quad \mathcal{A}_{\text{s,random}}(\mathfrak{J}, \lambda) \geq \mathcal{A}_{\text{random}}(\mathfrak{J}, \lambda). \quad (88)$$

The reverse statements can be proven by application of Lemma 17 in [7], which originally appeared in [22]: If a strong subspace transmission code for an AVQC  $\mathfrak{J}$  with error  $\lambda_l \in [0, 1]$  is given, then this directly implies the existence of an entanglement transmission code with error  $\hat{\lambda}_l = 1 - \frac{k_l+1}{k_l}(1 - \lambda_l - \frac{1}{k_l+1})$ , and of course if  $\liminf_{l \rightarrow \infty} \frac{1}{l} \log k_l = R > 0$  and  $\limsup_{l \rightarrow \infty} \lambda_l = \lambda$  then  $\limsup_{l \rightarrow \infty} \hat{\lambda}_l = \lambda$  and thus: If the r.h.s. below are nonzero, then

$$\mathcal{A}_{\text{det}}(\mathfrak{J}, \lambda) \geq \mathcal{A}_{\text{s,det}}(\mathfrak{J}, \lambda) \quad \text{and} \quad \mathcal{A}_{\text{random}}(\mathfrak{J}, \lambda) \geq \mathcal{A}_{\text{s,random}}(\mathfrak{J}, \lambda) \quad (89)$$

holds. But by equations (88), if the r.h.s. above are zero, then the l.h.s. are zero as well.  $\square$

*Acknowledgements.* This work was supported by the DFG via grant BO 1734/20-1 (H.B.) and by the BMBF via the grants 01BQ1050 and 16KIS0118 (H.B., J.N.).

## References

- [1] R. Ahlswede, “A Note on the Existence of the Weak Capacity for Channels with Arbitrarily Varying Channel Probability Functions and Its Relation to Shannon’s Zero Error Capacity” *The Annals of Mathematical Statistics*, Vol. 41, No. 3. (1970)
- [2] R. Ahlswede, “Elimination of Correlation in Random Codes for Arbitrarily Varying Channels”, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 44, 159-175 (1978)
- [3] R. Ahlswede, “Coloring Hypergraphs: A New Approach to Multi-user Source Coding-II”, *Journal of Combinatorics, Information & System Sciences* Vol. 5, No. 3, 220-268 (1980)

- [4] R. Ahlswede, "Arbitrarily Varying Channels with States Sequence Known to the Sender", *IEEE Trans. Inf. Th.* Vol. 32, 621-629, (1986)
- [5] R. Ahlswede, V. Blinovskiy, "Classical capacity of classical-quantum arbitrarily varying channels", *IEEE Trans. Inf. Theory*, Vol. 53, No. 2, 526-533.
- [6] R. Ahlswede, N. Cai, "Correlated sources help the transmission over AVC", *IEEE Trans. Inf. Th.*, Vol. 43, No. 4, 1254-1255 (1997)
- [7] R. Ahlswede, I. Bjelakovic, H. Boche, J. Nötzel "Quantum capacity under adversarial noise: arbitrarily varying quantum channels", *Comm. Math. Phys.*, Vol. 317, Iss. 1, 103-156 (2013)
- [8] R. Ahlswede, J. Wolfowitz, "The structure of capacity functions for compound channels", *Proc. of the Internat. Symposium on Probability and Information Theory at McMaster University, Canada*, 12-54, (1968)
- [9] R. Ahlswede, J. Wolfowitz, "The Capacity of a Channel with Arbitrarily Varying Channel Probability Functions and Binary Output Alphabet" *Z. Wahrscheinlichkeitstheorie verw. Geb.* 15, 186-194 (1970)
- [10] , H. Barnum, J. A. Smolin, B. M. Terhal, "Quantum capacity is properly defined without encodings", *Phys. Rev. A*, Vol. 58, No. 5, 3496-3501 (1998)
- [11] I. Bjelaković, H. Boche, "Classical Capacities of Averaged and Compound Quantum Channels", *IEEE Trans. Inf. Th.* Vol. 55, No. 7, 3360 - 3374 (2009)
- [12] I. Bjelakovic, H. Boche, G. Janßen, J. Nötzel, "Arbitrarily varying and compound classical-quantum channels and a note on quantum zero-error capacities", *Information Theory, Combinatorics, and Search Theory LNCS* Vol. 7777, 247-283 (2013)
- [13] I. Bjelaković, H. Boche, J. Nötzel, "Entanglement transmission and generation under channel uncertainty: Universal quantum channel coding", *Commun. Math. Phys.* 292, 55-97 (2009)
- [14] I. Bjelaković, H. Boche, J. Sommerfeld, "Secrecy Results for Compound Wiretap Channels", *Probl. Inf. Trans.* Vol. 49, Iss. 1, 73-98 (2013)
- [15] D. Blackwell, L. Breiman, A.J. Thomasian, "The capacities of certain channel classes under random coding", *Ann. Math. Stat.* 31, 558-567 (1960)
- [16] H. Boche, M. Cai, C. Deppe, "Classical-Quantum Arbitrarily Varying Wiretap Channel - A Capacity Formula with Ahlswede Dichotomy - Resources", *arXiv:1307.8007* (2013)
- [17] H. Boche, J. Nötzel, "Arbitrarily small amounts of correlation for arbitrarily varying quantum channels", *J. Math. Phys.* Vol. 54, 112202 (2013)
- [18] I. Csiszar, P. Narayan, "The Capacity of the Arbitrarily Varying Channel Revisited: Positivity, Constraints", *IEEE Trans. Inf. Th.* Vol. 34, No. 2, 181-193 (1989)
- [19] I. Devetak, "The private classical capacity and quantum capacity of a quantum channel", *IEEE Trans. Inf. Th.* 51, No.1, 44-55 (2005)
- [20] R. Duan, S. Severini, A. Winter, "Zero-error communication via quantum channels, non-commutative graphs and a quantum Lovász  $\theta$  function", *IEEE Trans. Inf. Theory*, 59(2):1164-1174, (2013)
- [21] T. Ericson, "Exponential Error Bounds for Random Codes in the Arbitrarily Varying Channel", *IEEE Trans. Inf. Th.* Vol. 31, No. 1, 42-48 (1985)
- [22] M. Horodecki, P. Horodecki, R. Horodecki, "General teleportation channel, singlet fraction, and quasidistillation", *Phys. Rev. A* 60, 18881898 (1999)

- [23] M. Horodecki, P. W. Shor, M. B. Ruskai, “General Entanglement Breaking Channels”, *Rev. Math. Phys.*, Vol. 15, 629-641 (2003)
- [24] M. Keyl, R. F. Werner, “How to Correct Small Quantum Errors”, *LNP* Vol. 611, 263-286 (2002)
- [25] J. Kiefer, J. Wolfowitz, “Channels with arbitrarily varying channel probability functions”, *Information and Control* Vol. 5, 44-54 (1962)
- [26] A.Yu. Kitaev, A.H. Shen, M.N. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics 47, American Mathematical Society, Providence, Rhode Island (2002)
- [27] D. Leung, G. Smith, “Continuity of quantum channel capacities”, *Commun. Math. Phys.* Vol. 292, 201-215, (2009)
- [28] J. Nötzel, “Quantum communication under channel uncertainty”, Ph.D. thesis, <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20121031-1100407-1-5> (2013)
- [29] Quantum information problem page of the ITP Hannover, <http://qig.itp.uni-hannover.de/qiproblems/11>
- [30] C. E. Shannon, “The zero error capacity of a noisy channel”. *IRE Trans. Information Theory* IT-2, 8-19 (1956)
- [31] B. Schumacher, “Sending entanglement through noisy quantum channels” *Phys. Rev. A* 54, 2614 (1996)
- [32] M. Wiese, H. Boche, “Strong Secrecy for Multiple Access Channels”, *Information Theory, Combinatorics, and Search Theory LNCS* Vol. 7777, 71-122, (2013)
- [33] W. Rudin, “*Principles of Mathematical Analysis*”, 3rd. edition, Mc Graw-Hill, Inc. 1976